

J. Perinat. Med.  
16 (1988) 273

## A prototype system for perinatal knowledge engineering using an artificial intelligence tool

Robert J. Sokol and Lawrence Chik

Department of Obstetrics and Gynecology, Wayne State University/Hutzel Hospital, Detroit, Michigan, U.S.A.

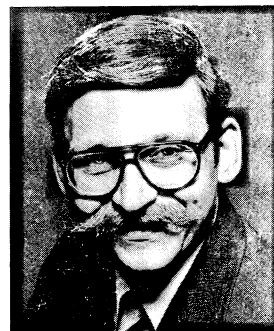
### 1 Introduction

Expert systems are generally regarded as programs that exhibit "intelligent," i.e., human-like attributes. These attributes typically would include the ability 1) to conduct free-form dialogue, 2) to provide explanations as to 'why' a response was produced when asked, 3) to accept new knowledge and retract erroneous or obsolete facts and 4) to resolve problems which were NOT explicitly pre-programmed. Expert systems using artificial intelligence (AI) programming tools are becoming more widely available for use in large corporate mainframe systems, as well as in microcomputing environments. Such systems are beginning to have impact in the business world. Some specialized applications have been investigated in medicine as well. The best known of the medical expert systems include the MYCIN [1] and the INTERNIST [2] systems. These systems are speculated to have involved 15 to 30 people years of development time. Because such systems have not been ported to laboratory scale or, preferably, departmental scale or personal computers, their impact on medical education and clinical medicine has been limited.

To the time of the First World Symposium on Computers in the Care of the Mother, Fetus and Newborn (March, 1987), there had been many medical application programs termed "expert systems" by their authors. Typically, these programs consist of diagnostic and/or treatment algorithms represented as explicit logical statements programmed in standard languages, including Basic and Fortran. Thus, for example, the current authors reported the development and utility of such a program for the diagnosis of normal and ab-

### Curriculum vitae

ROBERT J. SOKOL, M.D. is currently Professor and Chairman of the Department of Obstetrics and Gynecology at Wayne State University and Chief of the Department in the Detroit Medical Center/Hutzel Hospital. With about 450 publications, his major research interests include computer applications in perinatal medicine and the assessment of prenatal risks as determinants of infant outcome, particularly neurobehavioral development. Grant support is in the area of alcohol, as a perinatal risk.



normal labor progression 14 years ago [5]. Based upon the patient's parity, the time of onset of labor and the time, dilatation and station of each vaginal examination, that program, still in use, generates any of 53 appropriate diagnostic messages, singly or in combination. This type of application program can serve very useful purposes at the bedside for clinical care, educationally and as a research tool. Though other "expert systems" have been reported by perinatologists, to the best of our knowledge, the "intelligent" attributes of modern AI programming tools have not been used in significant applications in perinatal medicine. Yet, the potential for "an intelligence machine" remains conceptually promising, and ever more so with progressively more powerful departmental computers and workstations.

Our goal in the current study was to develop a useful perinatal knowledge base and inference engine for perinatal consultation at the level of the senior resident and perinatal fellow. This manuscript describes our initial experience in attempting to apply an artificial intelligence programming tool, 'PROLOG,' with a focus on the problems encountered in using the standard "bottom up" approach to building a knowledge base. Finally, we will describe a "top down" approach, i.e., global to the development of a perinatal knowledge base and inference engine and the resulting prototype system using a variety of system utilities in conjunction with PROLOG.

## 2 A "bottom up" approach

Inasmuch as the current investigators had no experience in developing artificial intelligence-based expert systems, we decided to employ the textbook approach and mimic the process described in the development of such systems in industry. Typically, there are two participants necessary for the development of such systems. The first is the "expert," i.e., the individual whose brain is to be "picked" for facts and rules for the knowledge base to be developed. In the current situation, this individual was termed **the perinatal expert**. The role of this individual was to identify a relevant area to specify specific topics and problems and to describe perinatal risks, appropriate medical interventions, including both workup and treatment, and to define expected perinatal outcomes, as well as to provide explanations for the various relationships included in the knowledge base.

In developing an expert system, the second necessary individual is **the knowledge engineer**. This individual's role includes first, developing an appropriate framework for facts and decision logic. Second, he/she must design a logic program for decision support and develop a program environment for both the interface with the user and the expert, i.e., for the consultative use of the system. Finally, the knowledge engineer must be concerned with the design of a natural language interface for both input and output from the system.

For development of our prototype system, a VAX 11/750 was available. The operating system is the University of California at Berkeley 4.1 BSD Unix [6]. After an evaluation of available AI program-

ming languages, we chose the University of New Hampshire NH Prolog. Following primarily the POPRAS [4] uniform perinatal record, we chose to experiment with a limited risk intervention outcome model, using the framework — risk, workup, treatment, outcome and explanation. An example of the framework and a single entry within this framework is shown in figure 1.

```

--- UNH PROLOG 1.3 ---
      (initialization)

WHAT IS THE TOPIC THAT YOU ARE INTERESTED IN ? I need information on iugr.

PLEASE WAIT, I AM LOOKING FOR SIMILAR TERMS.

SEARCH LIST:
* [IUGR]                                * [INTRAUTERINE, GROWTH, RETARDATION]
* [FETAL, GROWTH, RETARDATION]          * [GROWTH, RETARDATION]

** SEARCHING FRAMES **

PROBLEM TYPE: POOR OB HISTORY
SYMPTOM:      PREVIOUS SGA
WORKUP:       EARLY DETECTION BY SERIAL MCDONALD MEASUREMENTS OR
               ULTRASOUNDS, NSTS IN THIRD TRIMESTER
OUTCOME:      IUGR FETUS
REASON:       RISK TENDS TO RECUR BASED IN PART
               ON PERSISTENT CAUSES SUCH AS CHRONIC HYPERTENSION,
               SMOKING OR SUBSTANCE ABUSE

** SEARCHING TEXT **

INTRAUTERINE GROWTH RETARDATION AND CONGENITAL ANOMALIES ALSO CONTRIBUTE TO
NEONATAL MORTALITY. (5)

SPECIFIC ANTEPARTUM RISKS FOUND TO BE ASSOCIATED WITH NEONATAL MORBIDITY INCLUDE
MATERNAL CARDIOVASCULAR - RENAL COMPLICATIONS, IN PARTICULAR, PREGNANCY -
INDUCED HYPERTENSION, DIABETES MELLITUS, CLINICALLY SUSPECTED INTRAUTERINE
GROWTH RETARDATION, AND OTHER SEVERE MEDICAL / OBSTETRIC COMPLICATIONS.

FURTHERMORE, THE COMBINATION OF CHRONIC AND PREGNANCY - INDUCED HYPERTENSION,
IS ASSOCIATED WITH A CONSIDERABLE INCREASE IN RISK FOR INTRAUTERINE GROWTH
RETARDATION AND ABRUPTIO PLACENTA.

CYANOTIC HEART DISEASE IS A CLEAR RISK FOR INTRAUTERINE GROWTH RETARDATION. (25)

ALCOHOL IS A WELL ESTABLISHED TERATOGEN AND IS ASSOCIATED WITH INTRAUTERINE
GROWTH RETARDATION, AS WELL AS FETAL ALCOHOL SYNDROME AND ASSOCIATED MENTAL
RETARDATION. (30)

ACUTE FATTY LIVER OF PREGNANCY CONTINUES TO HAVE A HIGH MORTALITY RATE FOR THE
MOTHER. (37) STOMACH STAPLING, A TREATMENT FOR OBESITY WHICH HAS FALLEN OUT OF
FAVOR MAY SOMETIMES STILL BE ENCOUNTERED AND HAS BEEN ASSOCIATED DURING
PREGNANCY WITH AN INCREASED RISK OF INTRAUTERINE GROWTH RETARDATION AND WITH THE
RISK FOR MATERNAL HEPATIC NECROSIS. (38)

PERHAPS, THE MOST COMMONLY MISSED PROBLEM ASSOCIATED WITH THE UTERINE SIZE
BEING SMALLER THAN EXPECTED, IS INTRAUTERINE GROWTH RETARDATION, THE CLINICAL
INTERPRETATION BEING THAT THE PATIENT HAS NOT PROGRESSED AS FAR IN PREGNANCY AS
SHE HAD THOUGHT.
.....
DO YOU WANT TO CLEAR THE INPUT FOR RECURSION <Y/N>? y
DO YOU WANT A PHRASE LIST FROM THE OUTPUT FOR RECURSION <Y/N>? y

```

**Figure 1.** Edited printout for prototype perinatal consultation session. The user input follows the "?" and the blank lines between paragraphs in 'searching text' imply user options to continue. Numbers in parentheses indicate reference in on-line bibliography.

After, perhaps, 50 person hours of work, it "rapidly" became evident that this task was not going to be as easy as it first looked. In normal programming, a set of well-defined input/output specifications is passed to a programmer, who then implements those specifications in a set of program statements. In building a knowledge base, the "expert" must participate throughout the process. Furthermore, it became clear that the breadth of the task we had chosen, i. e., to develop a perinatal consultant was much too broad. Either, we would need to change our overall goal or our approach. We chose the latter route.

### 3 A "top down" approach

A straightforward description of the top down approach we envisioned was one which is more like doing a sculpture than building a house. It was readily apparent to us that there is a vast amount of information available, first, from the literature and second, from service statistics from perinatal data bases. We, therefore, focused our efforts on developing methods for the appropriate orderly retrieval of such information, using AI tools where relevant. We determined that for the prototype under development, there would be two primary sources of knowledge. The first of these was the service statistics for one year's experience on the Obstetric Service at Hutzell Hospital/Wayne State University. These statistics include the frequencies of antepartum and intrapartum risks and complications, as well as variables covering immediate neonatal outcome. The second source of "perinatal knowledge" was a manuscript we had recently prepared for a resident level textbook [3]. This manuscript covered, in broad general outline, the diagnostic methods and treatment of a full range of high risk pregnancy complications. The manuscript was 30 double-spaced typed pages and included 715 sentences in 106 paragraphs, excluding references.

Accessing information from the service statistics was straightforward and will not be described here in detail. Rather, we focus on the complex issue of extracting facts from a manuscript to build a perinatal knowledge base. Using the chapter as a basis and after trying several approaches, we came to the conclusion that the UNIX writing style analysis utility 'style' could be useful in retrieving some information from the manuscript by looking for the parts of speech.

### 4 Knowledge extraction from a manuscript: methods

For this study, two methods of extracting knowledge from the literature, here represented by a single chapter, broadly covering perinatal issues, were compared. We used the first method as our "gold standard" and termed this method "semi-handmade." The manuscript was processed into one sentence per line using UNIX utilities 'sed' and 'nroff' and parsed for the parts of speech sentence-by-sentence using 'style.' A Fortran program was written to compile the parts of speech to produce a list of key words and phrases (nouns and noun phrases). The resulting list was displayed on a CRT and edited by the perinatal expert, with the help of the knowledge engineer. This was relatively simple, inasmuch as the task was limited to eliminating words and phrases which, for the purpose at hand, did not identify important perinatal concepts.

The second method was termed "semiautomated." In this method, the manuscript was "spelled" first using a medical dictionary to validate the correct spelling, and then using a dictionary without medical words. As the output of the spelling program typically includes misspelled words and words unknown in the dictionary, the use of a nonmedical dictionary allowed us to simply identify medical terms and the corresponding sentences in the context.

The results of the "semihandmade" and "semiautomated" methods of identifying key terms are shown in table I. It can be seen that using the "semihandmade" approach, after editing, there were 664 key nouns and phrases representing perinatal concepts, i. e., risks, interventions and outcomes. In comparison, using the spell program, approximately half as many terms were identified.

Our next step was to examine and compare the success of our two methods of extracting information from the manuscript. The criterion function chosen a priori was the extraction of pertinent sentences representing medical knowledge from the manuscript. The results of this comparison are shown in the contingency table in table II. Using the gold standard "semihandmade" method, key sentences were taken as including greater than or equal to two key nouns and/or phrases. Of the 715 sentences in the manuscript, 30% (213) were identified as key by the experts. Of these 213, 169 were also identified by the semiautomated method using the spell check program (sensitivity = 79%).

**Table I.** Analysis of nouns and noun phrases by ‘style’ analysis (semihandmade) and analysis of medical terms by ‘spell’ (semiautomated) in manuscript

	Number	Uniq-sort	Edited
Semihandmade			
Key nouns	1673	815	241
Noun phrases	1780	1332	423
Total			664
Semiautomated			
Medical terms		312	

**Table II.** Correlation of 715 key sentences by semi-handmade method and semiautomated methods

	Semihandmade		
	Key sentences	Other sentences	
Semiautomated			
1 or more medical terms in sentences	169	146	315
no medical terms in sentences	44	356	400
	213	502	715

Sensitivity: 79%  
Specificity: 71%  
Positive productive value: 54%  
Negative productive value: 89%  
Chi-square: 153.3  
p < .001  
Adjusted R<sup>2</sup>: 39.8%

Twenty-one percent of the key sentences were not identified. Of 502 sentences not judged to contain key information by the experts, the semiautomated method, based on the lack of a medical word in the sentence, correctly identified 356 (specificity = 71%).

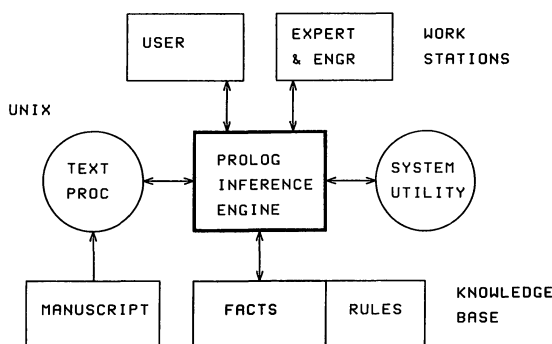
The semiautomated method chose 315 sentences of the 715 as important, based on the presence of at least one medical term. Of these 315, 169 were judged to be key by the experts for a positive predictive value of 54%. Similarly, of the 400 sentences that did not have a medical word, 356 were judged by the experts not to be key, for a negative predictive value of 89%. Calculated from chi-square and the marginals of the contingency table, as a measure of association, the adjusted R<sup>2</sup> was 39.8%.

Based on these results, we concluded that the spell program represented a useful step of selecting key sentences from this chapter manuscript. Considering that this manuscript was reasonably representative of other materials in the perinatal area, e. g., using very similar terms, we speculate that this approach may be useful for “noise disposal” in perinatal knowledge acquisition. However, the complexity of natural language will clearly necessitate better understanding of grammar rules and semantics for more precise information retrieval or “knowledge extraction”.

**5 A prototype system for knowledge engineering**

Having developed a potentially practicable method for knowledge extraction, we developed an overall concept of a prototype system for knowledge engineering. This is shown graphically in figure 2. Using text processing utilities from UNIX, information may be extracted from one or more manuscripts in a given perinatal area. Key sentences are then collected and coded into PROLOG compatible files as facts and rules, yielding a knowledge base. An inference engine, which could invoke system utilities, allows the expert and the knowledge engineer to interact with the system to ‘assert’ new facts and ‘retract’ erroneous and obsolete facts, as well as allowing the user to consult the knowledge base. In the prototype perinatal expert system being reported here, the knowledge base includes four components:

- 1) A similar term list. We have found this necessary to make it possible for the computer to identify all appropriate areas within the remainder of the knowledge base. Thus, for example, the computer must “know” that small-for-gestational age and intrauterine growth retardation, as well as fetal growth retardation, represents similar concepts.
- 2) Service statistics, as described above.
- 3) Key sentences as extracted by the methodology described above. These are expressed as PROLOG “facts” with appropriate syntax.
- 4) Finally, the full manuscript, along with its references, is included. This provides a simple solution to the “natural language interface problem” on the output side of the loop. The machine responds in natural language prepared by an expert.



**Figure 2.** A prototype system for knowledge engineering.

The inference engine in our prototype system includes preliminary rules and meta rules for an intelligent history taking program modeled after the POPRAS forms. In addition, portions of a recursive literature search program, using PROLOG, have been implemented.

An edited printout for a consultation with this prototype is shown in figure 1. To consult the system, the user might proceed as if he were using a medical book. Instead of using the table of contents or index, he/she might enter a key word or noun phrase of interest, e. g. iugr. Alternately, he/she might conduct a dialogue with the terminal and enter "I need information on intrauterine growth retardation." Using the parsing routine, the program identifies the key term, i. e., intrauterine growth retardation, and then proceeds serially through the frame structured knowledge base (risk, intervention and outcome) and the service statistics which include information on the frequency of diabetes mellitus and growth retardation and their relationship to perinatal outcome. Next, the key sentences are evaluated and appropriate segments of the entire manuscript, along with references, may be output under interactive control.

We are experimenting with various types of output and the order of output to determine what will most likely be of most use to the individual consulting the knowledge base. It may be noted that sometimes an in-depth search concerning a single subject is most appropriate, with an exhaustive search for everything in the knowledge base, e. g., concerning intrauterine growth retardation, prior to looking for hypertension or other related concepts. Alternatively, the user may wish to identify

only the intercepts, i. e., areas in the knowledge base in which IUGR is related to hypertension. Under interactive control, it is possible to branch from, say the first identification of IUGR in the knowledge base, to other key terms found in sentences mentioning IUGR, hypertension, or both. This latter approach allows for the rapid assessment of knowledge within a broad perinatal context.

At the end of run, the retrieved material can be analyzed to produce a new list of relevant topics for recursive search. The entire process emulates the response of an instructor to an inquisitive student.

## 6 Further steps

Based on our results to date, it is clear that considerably more work on the natural language interface for query input is necessary for free form dialogue. The relatively simple grammar rules of 'style' limit acceptable dialogue to relatively simple sentences only.

So far, we have only evaluated our semiautomatic "fact" extraction methodology on a single manuscript. Other manuscripts will need to be processed, using this methodology to obtain a better estimate of its utility. Also, considerably more work on the engine, including both the input/output interface, is needed to determine the most useful strategies for effective consultation, i. e., including interactive choice for depth versus breadth search paths.

## 7 Speculations and conclusions

Based on the experience gained over the past year and the results presented in this paper, it appears that a generalized approach to reducing the size of the immense task of perinatal knowledge engineering may be attainable. If so, it will undoubtedly be more efficient than hand building a multitude of individual knowledge bases of very limited breadth. Since the prototype system reported here functions in a computing environment with the capacity of a contemporary workstation, perhaps, practical educational and clinical applications might be possible. It is clear that an AI tool, such as PROLOG has the versatility to deal with many aspects of the list processing applications;

but it is not a necessary tool, in that other system utilities and conventional programming tools can be employed to accomplish the same goal. Based on the experience reported here, it is reasonable

to suggest that most elaborate expert system designs will require a combination on AI and conventional programming tools for design and development.

### Abstract

Though several perinatal expert systems are extant, the use of artificial intelligence has, as yet, had minimal impact in medical computing. In this evaluation of the potential of AI techniques in the development of a computer based "Perinatal Consultant," a "top down" approach to the development of a perinatal knowledge base was taken, using as a source for such a knowledge base a 30-page manuscript of a chapter concerning high risk pregnancy. The UNIX utility "style" was used to parse sentences and obtain key words and phrases, both as part of a natural language interface and to identify

key perinatal concepts. Compared with the "gold standard" of sentences containing key facts as chosen by the experts, a semiautomated method using a nonmedical speller to identify key words and phrases in context functioned with a sensitivity of 79%, i. e., approximately 8 in 10 key sentences were detected as the basis for PROLOG, rules and facts for the knowledge base. These encouraging results suggest that functional perinatal expert systems may well be expedited by using programming utilities in conjunction with AI tools and published literature.

**Keywords:** Artificial intelligence, clinical decision support, data base, expert systems, perinatology, PROLOG, UNIX.

### Zusammenfassung

#### Ein Systemprototyp in Form eines Werkzeugs der künstlichen Intelligenz für perinatales Ingenieurwesen in der Wissenschaft

Expertensysteme werden allgemein als Programme verstanden, die „intelligente“, d. h. menschenähnliche Attribute aufweisen. Diese Attribute beinhalten bezeichnenderweise die Fähigkeit,

- 1) freie Dialoge zu führen
- 2) bei Anfrage Erklärungen zu liefern, „warum“ eine Reaktion erfolgte,
- 3) neue Erkenntnisse zu übernehmen und fehlerhafte bzw. überholte Tatsachen zu verwenden und
- 4) Probleme zu lösen, die nicht explizit vorprogrammiert wurden.

Expertensysteme, die mit Werkzeugen der künstlichen Intelligenz arbeiten, verbreiten sich zunehmend in der Geschäftswelt, hatten jedoch, obwohl einige spezielle Anwendungen in der Medizin erforscht worden sind, bisher geringfügigen Einfluß auf die medizinische Auswertung.

Es wurden bisher eine Reihe von medizinischen Anwendungsprogrammen unter der Bezeichnung „Expertensysteme“ vorgestellt, jedoch nach den Kenntnissen der Autoren sind „intelligente“ Attribute moderner AI Programmierwerkzeuge bisher noch nicht bei wesentlichen Anwendungen in der perinatalen Medizin zur Anwendung gekommen. Das Ziel der vorliegenden Arbeit war es deshalb, eine brauchbare perinatologische Wissensbasis und eine Entscheidungshilfe für perinatale Beratung auf der Ebene eines Oberarztes und eines Assistenzarztes zu entwickeln.

Anfangs versuchten nur ein „Experte auf dem Gebiet der Perinatalogie“ und ein „Wissensingenieur“, ein klas-

sisches AI-System dadurch zu entwickeln, daß sie ein Gerüst einer Risikoauflbereitung, Behandlungsergebnisse und Erläuterungen aufstellten. Nach etwa 50 Arbeitsstunden an diesem Prototypsystem wurde deutlich, daß diese Aufgabe nicht so leicht zu lösen war, wie es zunächst erschien. Ein besonderes Problem bestand darin, daß während des Aufbaus der Wissensbasis der „Experte“ die ganze Zeit hindurch am Prozeß teilnehmen muß.

Zusätzlich wurde klar, daß die gestellte Aufgabe, nämlich Entwicklung eines Beraters für perinatologische Fragen, viel zu weit gefaßt war. Deshalb wurde ein anderer Weg beschritten.

Da umfangreiches Informationsmaterial sowohl aus der perinatalen Literatur als auch aus perinatologischen Datenbanken von seiten der Dienstleistungsstatistiken zur Verfügung steht, konzentrierten sich die Bemühungen auf die Entwicklung von Methoden zur zweckmäßigen und folgerichtigen Auffindung dieser Informationen, ggf. unter Verwendung der AI-Werkzeuge. Als Grundlage für das Prototypsystem wurden die Dienstleistungsstatistiken einer großen geburtshilflichen Klinik über ein Jahr und ein kürzlich erstellter Entwurf eines grundlegenden Lehrbuchkonzeptes über Hochrisikoschwangerschaften verwendet. Dieses Lehrbuchkonzept bestand aus etwa 30 zweizeilig gedruckten Seiten mit 715 Sätzen (ohne Literaturhinweise) in 106 Paragraphen.

Bezüglich der Problemstellung, „Wissen“ aus der Literatur herauszuziehen, d. h. aus einem bestimmten Kapitel, wurde ein „halbautomatisches“ System mit dem Wissen, das die Experten aus diesen entsprechenden Kapiteln herausgezogen haben, verglichen. Das a priori festgelegte Auswahlkriterium war das Auffinden der

Sätze im Lehrbuchkonzept, die das medizinische Wissen repräsentierten. Das halbautomatische System arbeitete auf einer rechnergestützten elektronischen „Buchstabenbasis“, indem es ein Wörterbuch ohne medizinische Fachwörter verwandte, was damit gleichzusetzen ist, daß die medizinischen Ausdrücke einfach in den entsprechenden Sätzen im Gesamttext ausgesucht wurden. Von den 750 Sätzen des Lehrbuchkonzeptes wurden 30% von den Fachkräften als Schlüsselsätze identifiziert. Von diesen 213 Sätzen wurden 169 (Sensitivität = 79%) von dem halbautomatischen System durch das Buchstabenvergleichs-Programm identifiziert. Entsprechend wurden 21% der Schlüsselsätze nicht gefunden. Von den 502 verbliebenen Sätzen, denen von den Fachkräften keine Schlüsselinformation zugeordnet worden war, hat das halbautomatische Verfahren, allein gestützt auf nicht-medizinische Wörter, 356 (Spezifität = 71%) richtig erkannt.

Auf der anderen Seite hat die halbautomatische Methode, ausgehend von wenigstens 1 medizinischen Ausdruck, von den 715 Sätzen 315 als richtig ausgewiesen. Von diesen 315 Sätzen wurden 169 von den Fachkräften eine Schlüsselrolle, was einem positiven Vorhersagewert von 54% entspricht, zugeordnet. Entsprechend wurde 356 Sätzen von den verbliebenen 400 Sätzen, die kein medizinisches Wort enthielten, von den Fachkräften keine Schlüsselrolle zuerkannt, was einem negativen Vorhersagewert von 89% entspricht. Dieses Ergebnis führt zu dem Schluß, daß das Buchstabenvergleichsprogramm einen hilfreichen Schritt bei der Auswahl von Schlüsselsätzen aus diesem in Kapitel unterteilten Lehrbuchkonzept darstellt.

In dem hier vorgestellten perinatologischen Prototypsystem wurden die nachfolgend aufgeführten 4 Komponenten als Wissensgrundlage eingeschlossen:

- 1) eine Liste ähnlicher Ausdrücke, über die es dem Computer ermöglicht wird, zugehörige Bereiche im verbleibenden Rest der Wissensgrundlage zu suchen,
- 2) Dienstleistungsstatistiken

3) Schlüsselsätze, die durch die oben beschriebene Methode herausgezogen, die in entsprechender Syntax als PROLOG-Tatsachen dargestellt wurden, und

4) schließlich das komplette Lehrbuchmanuskript mit seinen Referenzen, das bezüglich der Ausgabenseite der Kette eine einfache Lösung des Problems der Verbindung zur menschlichen Ausdrucksweise darstellt.

Die Maschine antwortet in einer vom Experten erstellten menschlichen Sprache. Die maschinelle Entscheidung innerhalb dieses Prototypsystems schließt vorläufige und hochbewertete Vorschriften für eine intelligente Übernahme der Krankengeschichte, aufbereitet gemäß PO-PRAS-Formblättern, ein. Darüber hinaus wurden Teile einer rekursiven Literatursuche unter dem Programm PROLOG implementiert. Wird dem System eine Beratungsanfrage gestellt, so wird diese vom System so interpretiert, als würde der Lehrer dem wißbegierigen Schüler antworten.

Es wurde dargelegt, daß das beschriebene System ein Prototyp ist, das noch wesentlicher Arbeiten bezüglich der Verbindung zur menschlichen Ausdrucksweise bei Anfragen und ebenfalls noch einer Entwicklung bezüglich der halbautomatischen Extraktion von Fakten aus der Literatur bedarf. Nichtsdestoweniger erscheint es, wie die hier vorgelegten Erfahrungen zeigen, vernünftig zu sein, daß ein derartiger Ansatz nicht nur möglich, sondern auch praktisch, erzieherisch und klinisch verwertbar ist. Ein AI-Werkzeug, wie PROLOG es darstellt, hat in vielfacher Hinsicht die Flexibilität zur Listenverarbeitung; aber es ist auf der anderen Seite kein notwendiges Werkzeug, das in andere Systembibliotheken und konventionelle Programme eingebaut werden müßte, um dieselben Ziele zu erfüllen.

Ausgehend von den hier berichteten Erfahrungen erscheint die Annahme als vernünftig, daß die weitaus größte Zahl von Entwürfen für Expertensysteme AI und konventionelle Programmierwerkzeuge sowohl beim Aufbau als auch bei der Weiterentwicklung Bedeutung haben sollte.

**Schlüsselwörter:** Datenbank, Expertensysteme, klinische Entscheidungshilfe, künstliche Intelligenz, Perinatalogie, PROLOG, UNIX.

## Résumé

### Prototyp pour l'organisation des connaissances perinatales utilisant un instrument d'intelligence artificielle

On considère généralement les systèmes experts comme des programmes qui paraissent « intelligents », c'est-à-dire qu'ils possèdent des attributs humains. Ces attributs comprennent typiquement la capacité 1) de conduire un dialogue de forme libre, 2) de fournir des explications telles que « pourquoi » une réponse a été effectuée lorsqu'elle a été demandée, 3) d'accepter de nouvelles connaissances et de désavouer des faits erronés ou obsolètes, et 4) de résoudre ce qui n'a pas été explicitement pré-programmé. Des systèmes experts utilisant des outils de programmation d'intelligence artificielle (IA) deviennent plus largement disponibles dans le monde du

travail, mais, bien que certaines applications spécialisées aient été explorées dans le champ de la médecine, ils n'ont pour l'instant qu'un impact minime en informatique médicale.

De nombreux programmes d'application médicale qualifiés de « systèmes experts » par leurs promoteurs ont été proposés, mais à la connaissance des auteurs, les attributs « intelligents » des outils de programmation de l'IA moderne n'ont pas encore été utilisés dans des applications significatives en médecine périnatale. Le but de cette étude a été de développer une base de connaissances périnatales utiles et une machine déductive pour la consultation périnatale du niveau de l'interne expérimenté (senior resident) et du périnatal fellow.

Dans un premier temps, «un expert en périnatalogie» et «un ingénieur de la connaissance» se sont efforcés de développer un système «classique» d'IA, à l'aide d'un cadre d'élaboration de risque, de traitement de l'évolution et d'explication. Après environ 50 heures de travail sur ce système prototype, il est devenu évident que la tâche ne serait pas aussi aisée que ce qu'elle apparaissait de prime abord. Un problème particulier a été le suivant: l'expert doit participer tout au long du processus d'élaboration de la base de connaissance. En outre il est apparu clairement que l'ampleur de la tâche choisie, c'est-à-dire mettre au point un consultant en périnatalogie, était bien trop vaste. C'est pourquoi on a adopté une approche alternative.

Puisqu'il existe une grande quantité d'informations disponibles, en provenance de la littérature périnatale et des statistiques des services tirées des bases de données périnatales, on a centré les efforts sur la réalisation de méthodes de restitution appropriée et classé de telles informations en utilisant les outils d'IA où ils sont appropriés. Pour le système prototype, on a utilisé les statistiques d'une année d'activité d'un grand service d'obstétrique et un manuscrit récent rédigé pour un manuel sur les grossesses à haut risque s'adressant aux résidents. Le manuscrit était de 30 pages dactylographiées à double interligne et comprenait 715 phrases réparties en 106 paragraphes, à l'exception des références.

En se focalisant sur le problème de l'extraction «des connaissances» de la littérature, représentée par un unique chapitre, on a comparé un système «semi-automate» aux connaissances extraites de ce chapitre par les experts. Le critère choisi *a priori* était l'extraction de phrases pertinentes représentant les connaissances médicales du manuscrit. Le système semi-automate fonctionne au moyen d'un «alphabet» informatisé, utilisant un dictionnaire sans termes médicaux qui procure une méthode d'identification simple des termes médicaux au sein des phrases correspondantes au contexte. Parmi les 715 phrases du manuscrit, 30% ont été identifiées comme des critères par les experts. Parmi ces 213 phrases, 169 ont été identifiées par la méthode semi-automatique utilisant le programme de vérification alphabétique (sensibilité = 79%). Vingt et un pour cent des phrases critères n'ont pas été identifiées. Sur les 502 phrases qui n'ont pas été jugées contenir des informations critères par les experts, la méthode semi-automatique, fondée sur l'absence de termes médicaux dans la phrase, en a identifié correctement 356 (spécificité = 71%). La méthode semi-automatique reconnaît 315 phrases parmi les 715 comme importantes, en se fondant sur la présence

d'au moins un terme médical. Parmi ces 315 phrases, les experts en ont jugé 169 comme étant des critères, avec une valeur prédictive positive de 54%. De la même façon, parmi les 400 phrases qui ne contiennent pas de termes médicaux, les experts ont jugé que 356 n'étaient pas des critères, avec une valeur prédictive négative de 89%. En se fondant sur ces résultats, on a conclu que le programme alphabétique représentait une étape utile pour sélectionner les phrases critères de ce manuscrit.

Dans le système périnatal prototype rapporté ici, la base de connaissance comprend quatre composantes:

- 1) une liste des termes similaires, nécessaire pour que l'ordinateur puisse identifier toutes les zones appropriées au sein du reste de la base de connaissances;
- 2) des statistiques du service;
- 3) des phrases critères, extraites par la méthodologie décrite ci-dessous, exprimées en tant que «faits» PROLOG avec une syntaxe adéquate, et
- 4) enfin, la totalité du manuscrit, y compris ses références, ce qui assure une solution simple au «problème de l'interface avec le langage naturel» au niveau de la sortie de la boucle. La machine répond en langage naturel qui a été préparé par un expert. L'instrument déductif de ce système prototype comprend des règles préliminaires et des règles pour une histoire intelligente prenant un programme présenté après les formes PO-PRAS. En outre, des fragments de programme de recherche de littérature récursive, utilisant le PROLOG ont été ajoutés. Utilisé comme consultation, le processus dans son ensemble utilisant ce prototype imite la réponse d'un instructeur à un étudiant inquisiteur.

On insiste sur le fait qu'il s'agit d'un système prototype sur lequel un travail plus important est nécessaire tant au niveau de l'interface avec le langage naturel au niveau de l'entrée de questionnement, qu'au niveau de l'extraction au sein de la littérature de «faits» d'évaluation semi-automatique. Néanmoins, en se fondant sur l'expérience rapportée ici, il est raisonnable de suggérer que l'utilisation d'une telle approche rend bien possible des applications pratiques, éducationnelles et cliniques. Un outil d'IA, tel que le PROLOG, a une grande souplesse d'emploi pour coïncider avec les nombreux aspects des applications des processus de listing; mais ce n'est pas un outil nécessaire en ce sens que d'autres systèmes et d'autres outils de programmation conventionnelle peuvent être utilisés pour accomplir le même objectif. En se fondant sur l'expérience rapportée ici, il est raisonnable de suggérer que de nombreux types de systèmes experts nécessiteront l'association de l'IA et d'outils de programmation conventionnelle pour leur conception et leur développement.

**Mots-clés:** Aide à la décision clinique, base de données, intelligence artificielle, périnatalogie, PROLOG, système expert, UNIX.



## References

- [1] BARR A, EA FEIGENBAUM: The Handbook of Artificial Intelligence, Vol. II. HeurisTech Press, Stanford, CA, (1982) 184
- [2] BARR A, EA FEIGENBAUM: The Handbook of Artificial Intelligence, Vol. II. HeurisTech Press, Stanford, CA, (1982) 197
- [3] BHATIA R, RJ SOKOL, M PERNOLL: Detection of High Risk Pregnancy. In: PERNOLL M, BENSON RC: Current Obstetric and Gynecologic Diagnosis and Treatment, 6th Edition. Lange, Los Altos, CA, (1987) 246–254
- [4] HOBEL CJ: Development of POPRAS (Problem Oriented Perinatal Risk Assessment System). In: BAHR JP (ed): Proceedings First National Conference on Perinatal Data Systems, 1980. Wisconsin Assoc for Perinatal Care, Medical College of Wisconsin
- [5] SOKOL RJ, RS NUSSBAUM, L CHIK, M ROSEN: Computer Diagnosis of Labor Progression: I. Development of an On-Line Interactive Digital Computer Program for the Diagnosis of Normal and Abnormal Cervical Dilatation Patterns. J Reprod Med 11 (1973) 149
- [6] Unix Programmer's Manual, 4.1 BSD User Document, Computer Science Division, Department of Electrical Engineering and Computer Science, Univ. of California, Berkeley, CA, 1981

Robert J. Sokol, M.D.  
Department of Ob/Gyn  
Hutzel Hospital  
4707 St. Antoine  
Detroit, MI 48201, U. S. A.